

# SEQ SEARCH SUMMARY

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 16, 2003, 15:48:55 ; Search time 55475 Seconds  
(without alignments)  
11813.830 Million cell updates/sec

Title: US-09-830-160-1  
Perfect score: 16020  
Sequence: 1 agatctcggtccgccagtgcc.....caccggccgggagagatct 16020

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl:\*  
1: gb\_ba:\*  
2: gb\_htg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sts:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vi:\*  
15: em\_ba:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_mu:\*  
20: em\_om:\*  
21: em\_or:\*  
22: em\_ov:\*  
23: em\_pat:\*  
24: em\_ph:\*  
25: em\_pl:\*  
26: em\_ro:\*  
27: em\_sts:\*

28: em\_un:\*  
 29: em\_vi:\*  
 30: em\_htg\_hum:\*  
 31: em\_htg\_inv:\*  
 32: em\_htg\_other:\*  
 33: em\_htg\_mus:\*  
 34: em\_htg\_pln:\*  
 35: em\_htg\_rod:\*  
 36: em\_htg\_mam:\*  
 37: em\_htg\_vrt:\*  
 38: em\_sy:\*  
 39: em\_htgo\_hum:\*  
 40: em\_htgo\_mus:\*  
 41: em\_htgo\_other:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

| Result |       | %       |        | Query  |    | Match Length DB |  | ID       | Description |  |
|--------|-------|---------|--------|--------|----|-----------------|--|----------|-------------|--|
| No.    | Score | Match   | Length | DB     | ID |                 |  |          |             |  |
| c      | 1     | 15741.6 | 98.3   | 16065  | 1  | AF187532        |  | AF187532 | Streptomy   |  |
| c      | 2     | 1847    | 11.5   | 14807  | 1  | AF264025        |  | AF264025 | Streptomy   |  |
| c      | 3     | 1085.4  | 6.8    | 25883  | 1  | AB008466        |  | AB008466 | Streptomy   |  |
| c      | 4     | 882     | 5.5    | 9523   | 1  | AF237894S1      |  | AF237894 | Streptomy   |  |
| c      | 5     | 848.2   | 5.3    | 15052  | 1  | AF055579        |  | AF055579 | Streptomy   |  |
|        | 6     | 760.6   | 4.7    | 5084   | 1  | SGSTRELMB       |  | X62567   | S.griseus s |  |
| c      | 7     | 689.8   | 4.3    | 18977  | 1  | SAR7932         |  | AJ007932 | Streptomy   |  |
| c      | 8     | 686.2   | 4.3    | 25459  | 1  | SGAJ6985        |  | AJ006985 | Streptomy   |  |
| c      | 9     | 665.4   | 4.2    | 2311   | 1  | AF355468        |  | AF355468 | Saccharop   |  |
| c      | 10    | 665.4   | 4.2    | 2363   | 6  | AX089466        |  | AX089466 | Sequence    |  |
| c      | 11    | 661.4   | 4.1    | 2310   | 6  | AR165019        |  | AR165019 | Sequence    |  |
| c      | 12    | 661.4   | 4.1    | 2310   | 6  | AR281867        |  | AR281867 | Sequence    |  |
| c      | 13    | 661.4   | 4.1    | 2310   | 6  | BD137650        |  | BD137650 | Biosynthe   |  |
| c      | 14    | 661.2   | 4.1    | 3500   | 1  | SERORF1         |  | L37354   | Saccharopol |  |
|        | 15    | 656.6   | 4.1    | 4133   | 1  | AB054887        |  | AB054887 | Streptomy   |  |
| c      | 16    | 646.6   | 4.0    | 22999  | 1  | AB088119        |  | AB088119 | Streptomy   |  |
| c      | 17    | 636.6   | 4.0    | 56917  | 1  | AME16952        |  | Y16952   | Amycolatops |  |
| c      | 18    | 607.8   | 3.8    | 37941  | 1  | AOPCZA361       |  | AJ223998 | Amycolato   |  |
| c      | 19    | 596     | 3.7    | 6085   | 6  | A60304          |  | A60304   | Sequence 4  |  |
| c      | 20    | 596     | 3.7    | 6085   | 6  | AR144763        |  | AR144763 | Sequence    |  |
| c      | 21    | 580.8   | 3.6    | 37116  | 6  | AX574199        |  | AX574199 | Sequence    |  |
|        | 22    | 580.8   | 3.6    | 109519 | 6  | AX195929        |  | AX195929 | Sequence    |  |
| c      | 23    | 579.8   | 3.6    | 2313   | 1  | STMDNRI         |  | M80237   | S.peucetius |  |
|        | 24    | 579.6   | 3.6    | 38506  | 6  | AR271680        |  | AR271680 | Sequence    |  |
|        | 25    | 579.6   | 3.6    | 38506  | 6  | AR277664        |  | AR277664 | Sequence    |  |
|        | 26    | 578     | 3.6    | 12441  | 1  | AF079762        |  | AF079762 | Streptomy.  |  |
| c      | 27    | 577.8   | 3.6    | 1119   | 6  | AX195932        |  | AX195932 | Sequence    |  |
| c      | 28    | 577.8   | 3.6    | 1122   | 6  | AX574132        |  | AX574132 | Sequence    |  |
| c      | 29    | 577.8   | 3.6    | 1128   | 6  | AX574130        |  | AX574130 | Sequence    |  |
| c      | 30    | 577.8   | 3.6    | 10035  | 6  | AX574202        |  | AX574202 | Sequence    |  |
| c      | 31    | 577.8   | 3.6    | 11115  | 6  | AX205014        |  | AX205014 | Sequence    |  |
| c      | 32    | 577.8   | 3.6    | 47981  | 1  | AF263245        |  | AF263245 | Micromono   |  |
| c      | 33    | 577.8   | 3.6    | 47981  | 6  | AX112026        |  | AX112026 | Sequence    |  |

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 16, 2003, 15:48:55 ; Search time 3546 Seconds  
(without alignments)  
12195.431 Million cell updates/sec

Title: US-09-830-160-1  
Perfect score: 16020  
Sequence: 1 agatctcggtccgccagtgcc.....caccggccgggagagatct 16020

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_19Jun03:\*

- 1: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:\*
- 2: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:\*
- 3: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:\*
- 4: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:\*
- 5: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:\*
- 6: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:\*
- 7: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:\*
- 8: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:\*
- 9: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:\*
- 10: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:\*
- 11: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:\*
- 12: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:\*
- 13: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:\*
- 14: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:\*
- 15: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:\*
- 16: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:\*
- 17: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:\*
- 18: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:\*
- 19: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:\*
- 20: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:\*
- 21: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:\*
- 22: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:\*
- 23: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:\*
- 24: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:\*
- 25: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2003.DAT:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,

and is derived by analysis of the total score distribution.

# SUMMARIES

| Result |     |       | %     |        |       |          | DB | ID | Description        |
|--------|-----|-------|-------|--------|-------|----------|----|----|--------------------|
|        | No. | Score | Match | Length | Query | DB       |    |    |                    |
|        | 1   | 15920 | 99.4  | 16020  | 21    | AAA39283 |    |    | Streptomyces nogal |
| c      | 2   | 1847  | 11.5  | 14806  | 22    | AAD03809 |    |    | Streptomyces galil |
| c      | 3   | 665.4 | 4.2   | 2363   | 22    | AAF88340 |    |    | S. spinosa polyket |
| c      | 4   | 661.4 | 4.1   | 2310   | 20    | AAZ21502 |    |    | Saccharopolyspora  |
| c      | 5   | 596   | 3.7   | 6085   | 18    | AAT70153 |    |    | S.longisporoflavus |
| c      | 6   | 580.8 | 3.6   | 37116  | 25    | ABZ66810 |    |    | Orthosomycin biosy |
|        | 7   | 580.8 | 3.6   | 109519 | 22    | AAS08693 |    |    | Micromonospora DNA |
|        | 8   | 579.6 | 3.6   | 38506  | 21    | AAA75633 |    |    | Nucleotide sequenc |
|        | 9   | 579.6 | 3.6   | 38506  | 21    | AAZ56001 |    |    | Recombinant cosmid |
|        | 10  | 579.6 | 3.6   | 38506  | 25    | ABS56090 |    |    | S. venezuelae DNA  |
|        | 11  | 578   | 3.6   | 12441  | 21    | AAZ87284 |    |    | S. venezuelae deso |
|        | 12  | 578   | 3.6   | 13613  | 24    | AAD39043 |    |    | Streptomyces venez |
| c      | 13  | 577.8 | 3.6   | 1122   | 25    | ABZ66775 |    |    | Orthosomycin biosy |
| c      | 14  | 577.8 | 3.6   | 1128   | 25    | ABZ66774 |    |    | Orthosomycin biosy |
| c      | 15  | 577.8 | 3.6   | 10035  | 25    | ABZ66813 |    |    | Orthosomycin biosy |
| c      | 16  | 577.8 | 3.6   | 11115  | 23    | ABL50562 |    |    | Micromonospora car |
| c      | 17  | 577.8 | 3.6   | 47981  | 22    | AAF30757 |    |    | Micromonospora meg |
|        | 18  | 561.6 | 3.5   | 1281   | 21    | AAZ87292 |    |    | S. venezuelae deso |
|        | 19  | 561.6 | 3.5   | 1281   | 24    | AAD39050 |    |    | Streptomyces venez |
| c      | 20  | 530.2 | 3.3   | 3412   | 20    | AAX25772 |    |    | S.erythraea erythr |
|        | 21  | 522.2 | 3.3   | 3756   | 18    | AAT72684 |    |    | Sugar biosynthesis |
| c      | 22  | 518.6 | 3.2   | 1600   | 24    | AAS18441 |    |    | Contig 88 DNA enco |
| c      | 23  | 516.2 | 3.2   | 42000  | 21    | AAA63349 |    |    | Streptomyces globi |
| c      | 24  | 516.2 | 3.2   | 63164  | 21    | AAA63348 |    |    | Streptomyces globi |
|        | 25  | 510.4 | 3.2   | 6093   | 20    | AAX25775 |    |    | S.erythraea oleand |
|        | 26  | 510.4 | 3.2   | 50937  | 21    | AAA09469 |    |    | Streptococcus olea |
| c      | 27  | 486.8 | 3.0   | 6854   | 19    | AAT76903 |    |    | S. glaucescens Pst |
|        | 28  | 485.6 | 3.0   | 1476   | 24    | AAS18443 |    |    | Contig 122 DNA enc |
| c      | 29  | 479.4 | 3.0   | 1266   | 20    | AAX25773 |    |    | S.erythraea erythr |
| c      | 30  | 475.2 | 3.0   | 3035   | 21    | AAA63462 |    |    | Streptomyces globi |
| c      | 31  | 474   | 3.0   | 3291   | 21    | AAZ56004 |    |    | Contig 003 from co |
|        | 32  | 473.6 | 3.0   | 14806  | 22    | AAD03809 |    |    | Streptomyces galil |
| c      | 33  | 472.4 | 2.9   | 3292   | 21    | AAA75636 |    |    | Nucleotide sequenc |
| c      | 34  | 472.4 | 2.9   | 3292   | 25    | ABS56093 |    |    | Contig 003 from co |
|        | 35  | 470.8 | 2.9   | 13613  | 21    | AAZ87319 |    |    | S. venezuelae deso |
| c      | 36  | 469.2 | 2.9   | 9994   | 22    | AAC85191 |    |    | S. avermitilis 10  |
|        | 37  | 464.4 | 2.9   | 1272   | 20    | AAZ21503 |    |    | Saccharopolyspora  |
|        | 38  | 464.2 | 2.9   | 19016  | 25    | AAD47222 |    |    | Streptococcus sp.  |
| c      | 39  | 460.6 | 2.9   | 1014   | 21    | AAZ87289 |    |    | S. venezuelae deso |
| c      | 40  | 460.6 | 2.9   | 1014   | 24    | AAD39047 |    |    | Streptomyces venez |
| c      | 41  | 458.8 | 2.9   | 990    | 22    | AAF88341 |    |    | S. spinosa DNA fra |
|        | 42  | 452.8 | 2.8   | 9994   | 22    | AAC85191 |    |    | S. avermitilis 10  |
|        | 43  | 437.4 | 2.7   | 47981  | 22    | AAF30757 |    |    | Micromonospora meg |
|        | 44  | 432.2 | 2.7   | 3291   | 21    | AAZ56004 |    |    | Contig 003 from co |
|        | 45  | 432.2 | 2.7   | 3292   | 21    | AAA75636 |    |    | Nucleotide sequenc |

# ALIGNMENTS

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 16, 2003, 15:52:15 ; Search time 754 Seconds  
(without alignments)  
9377.928 Million cell updates/sec

Title: US-09-830-160-1  
Perfect score: 16020  
Sequence: 1 agatctcgtccgccagtgcc.....cacccggccgggagagatct 16020

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_NA:\*  
1: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq:\*  
2: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq:\*  
3: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq:\*  
4: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq:\*  
5: /cgn2\_6/ptodata/2/ina/PCTUS\_COMB.seq:\*  
6: /cgn2\_6/ptodata/2/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

|        |       | %     |        |       |    |                   |                   |
|--------|-------|-------|--------|-------|----|-------------------|-------------------|
| Result |       | Query |        |       |    |                   |                   |
| No.    | Score | Match | Length | DB    | ID | Description       |                   |
| c      | 1     | 661.4 | 4.1    | 2310  | 3  | US-09-036-987A-25 | Sequence 25, Appl |
| c      | 2     | 661.4 | 4.1    | 2310  | 3  | US-09-370-700-25  | Sequence 25, Appl |
| c      | 3     | 661.4 | 4.1    | 2310  | 4  | US-09-603-207-25  | Sequence 25, Appl |
| c      | 4     | 596   | 3.7    | 6085  | 3  | US-09-029-603-4   | Sequence 4, Appli |
|        | 5     | 579.6 | 3.6    | 38506 | 3  | US-09-320-878-19  | Sequence 19, Appl |
|        | 6     | 579.6 | 3.6    | 38506 | 4  | US-09-141-908-1   | Sequence 1, Appli |
|        | 7     | 579.6 | 3.6    | 38506 | 4  | US-09-657-440-19  | Sequence 19, Appl |
|        | 8     | 578   | 3.6    | 13613 | 3  | US-09-105-537-3   | Sequence 3, Appli |
| c      | 9     | 577.8 | 3.6    | 47981 | 4  | US-09-679-279-1   | Sequence 1, Appli |
|        | 10    | 561.6 | 3.5    | 1281  | 3  | US-09-105-537-19  | Sequence 19, Appl |
|        | 11    | 522.2 | 3.3    | 3756  | 2  | US-08-576-626A-1  | Sequence 1, Appli |
| c      | 12    | 518.6 | 3.2    | 1600  | 4  | US-09-434-288-10  | Sequence 10, Appl |

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 16, 2003, 15:52:15 ; Search time 4192 Seconds  
(without alignments)  
12489.655 Million cell updates/sec

Title: US-09-830-160-1  
Perfect score: 16020  
Sequence: 1 agatctcgtccgccagtgcc.....caccggccgggagagatct 16020  
Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0  
Searched: 2169961 seqs, 1634102185 residues  
Total number of hits satisfying chosen parameters: 4339922

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications\_NA:\*

- 1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq:\*
- 10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq:\*
- 12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*
- 13: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq:\*
- 14: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq:\*
- 15: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*
- 16: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*
- 17: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | % Query |              |       |    |                  |  | Description       |
|------------|---------|--------------|-------|----|------------------|--|-------------------|
|            | Score   | Match Length | DB    | ID |                  |  |                   |
| 1          | 579.6   | 3.6          | 38506 | 11 | US-09-793-708-19 |  | Sequence 19, Appl |
| 2          | 579.6   | 3.6          | 38506 | 12 | US-10-201-365-1  |  | Sequence 1, Appli |
| 3          | 579.6   | 3.6          | 38506 | 12 | US-10-160-539-19 |  | Sequence 19, Appl |
| 4          | 578     | 3.6          | 12441 | 11 | US-09-988-384B-3 |  | Sequence 3, Appli |
| 5          | 578     | 3.6          | 13613 | 10 | US-09-861-289-3  |  | Sequence 3, Appli |